AAAI 2016 Doctoral Consortium Application

September 22, 2015

Personal Information

Full Name	Benjamin Bengfort
Mailing Address	73 Bryant St NW Washington DC 20001
Email Address	bengfort@cs.umd.edu
Phone Number	(701) 680-3095
Affiliation	Department of Computer Science University of Maryland United States of America
Home Page	`http://cs.umd.edu/~bengfort/`
Gender	Male
Citizenship	United States of America

Academic Information

Expected Graduation	May, 2017
Thesis Advisor	Michael T. Cox 4035 Col. Glenn Highway Suite 200 Beavercreek, OH 45431
Advisor Email	michael.cox@wright.edu
AAAI Mentors	Dr. David Aha Dr. Sergei Nirenburg
Keywords	Case Based Reasoning Knowledge Goals Goal Trajectories Question/Answer Systems

Guiding Goal Trajectories During Interactive Knowledge-Goal Reasoning

Benjamin Bengfort
Department of Computer Science
University of Maryland
bengfort@cs.umd.edu

Michael T. Cox
UMD Institute for Advanced Computer Studies
Wright State Research Institute
michael.cox@wright.edu

Motivation

We claim that humans solve complex questions through *investigation*, breaking down harder questions into simpler sub-questions with more easily obtained solutions. When answering questions, humans also take into account context and approach new problems by leveraging techniques that have worked in the past. Underneath every question is a knowledge goal that represents the desire for answers. A system that models how humans investigate questions will therefore *deconstruct* complex knowledge goals, *integrate* relevant knowledge, and *reuse* methods of investigation.

Investigation can therefore be represented as a plan to solve the larger knowledge goal, and if the tasks or sub-goals of knowledge acquisition plans are simple knowledge goals that can be solved via structured data retrieval, then a system can be said to solve complex knowledge goals through a planning process that involves the purposeful combination of simpler knowledge goals. However, due to the exploratory nature of investigation, the original knowledge goal will change, usually through a series of intermediate goals creating a *goal trajectory*. We propose that a cognitive system can augment or replicate the investigative process through a study of goal trajectories and complex knowledge goal solutions.

Knowledge Goals

A *knowledge goal* represents the purposeful need to acquire information in order to fill in gaps of world knowledge for a reasoning entity or to extend the database of a computational understanding system (Ram and Hunter 1991). For humans, knowledge goals are most easily represented as questions, and current research on dialog-driven question and answer systems focuses on the semantic parsing of a natural language question to a structured database query (Yahya et al. 2012) or lambda calculus representation (Berant et al. 2013). These parsing approaches are making headway in the solution of simple knowledge goals, where the primary task is a retrieval from some structured knowledge base; especially goals that ask who, what, when, or where.

This approach, however, cannot solve complex knowledge goals including aggregations, opinions (recommendations), or explanations (why or how questions). As a result, even though information retrieval systems and search have made information easily accessible, there has been the growth of community-driven question sites such as Quora (Wang et al. 2013) to connect users to the more nuanced answers they are looking for. Instead, by taking a goal-driven approach to question and answer systems, we believe a cognitive system can take advantage of a hierarchy of goals and solutions to augment the human investigative processes in an interactive fashion.

Representation

The representation of a knowledge goal can be described by its objective: the type of the expected response (concept specification) and how the information will be used (task specification) (Ram 1991). To create a plan to solve a knowledge goal, a computational system must be able to represent knowledge goals in a structured manner, and identify and parse the structure from a natural language query. We extend this representation to leverage the additional information available in an interactive knowledge-goal reasoning system. Knowledge goals are composed of three primary features: a set of *concepts*, a *task*, and a specific *context*.

1. **Concept**: The knowledge goal is relevant to a specific set of structured semantic concepts, both entities and predicates, which are either directly specified or implied. The concept of the question identifies the domain of the query plan as well as the expected response.
2. **Task**: The task indicates the purpose of the knowledge goal and is usually implied in a natural language query and related to the motivation of the questioner. For example, in the case of simple knowledge goals the task may be search or database queries; for more complex knowledge goals the task may be explanation or knowledge generation.
3. **Context**: Knowledge goals are not specified in isolation, but are directly related to the questioner and as such must be considered in relation to their context. Contextual information includes temporal, geographic, hysteretic, and preference information and is essential for plan refinement and to resolve ambiguity.

Interactive Knowledge Goal Reasoning

Our proposed approach for complex knowledge goal reasoning is a *case-based reasoning* (CBR) (Kolodner 1993; Lopez De Mantaras et al. 2005) system which reuses past experience in an interactive fashion. Interactive CBR operates similarly to conversational case-based reasoning systems, which incrementally elicits a target problem through an interactive dialog with the user, attempting to minimize the number of questions before a solution is reached (Aha, McSherry, and Yang 2005). In order to provide an adaptable, investigative system, the methodology we are exploring guides the user in a finite length interactive dialogue, removing the requirement to minimize session length in order to facilitate an ongoing discovery process. Additionally, the system itself is a learning agent with the goal of predicting future knowledge goals, and acquiring the information in advance in order to provide specific guidance to the user.

Goal Trajectories

In interactive knowledge-goal reasoning, a complex knowledge goal is broken down into a hierarchical plan involving simpler sub-goals and tasks in order to augment the investigative process of a user. During an investigative dialogue, the user selects the next step of plans proposed by the reasoner, and continues to chain sub-goals together to work towards a larger solution. During this process, the plan is adaptable and subject to change as the user proceeds in selecting and executing the simpler goals. Goals, like plans, are also subject to change and can be transformed (Cox and Veloso 1998), therefore we propose that the original knowledge goal itself is also subject to change; and that as knowledge goals change during interactive reasoning, the path that led to the solution of new knowledge goals from the original can be represented by a *goal trajectory*.

Goal trajectories can be influenced either through the direct interactive manipulation of goals (Cox and Zhang 2007); via other users in the system issuing similar queries that provide the basis for recommending new goals; or by the monitoring of new information or relevant data that has been added to the knowledge base. An investigative dialogue can be therefore seen as a planning problem where knowledge goals are not static and must be responsive to goal changes. We believe that a system can leverage goal change to provide guidance by proposing medium steps towards a series of predicted goals. This guidance will accelerate the user who is likely to take short steps toward a goal, yet not provide uncanny or mystifying advice by proposing longer, unintuitive steps.

Plan of Study

At the time of this submission, I have created an online, user-driven question and answer system called *Kyudo* with the purpose of generating a casebase of goal trajectories. It currently contains a small demonstration case base of questions and responses in a concierge domain (tourists asking for travel advice in Washington D.C) created by fellow graduate students. The casebase acquisition and plans for interactive knowledge goal reasoning using this system were presented at the Goal Reasoning Workshop in the *Advances in Cognitive Systems 2015* conference (Bengfort and Cox 2015). From the feedback at this workshop, it became clear that describing knowledge goal change in an investigative system could be of interest to the goal reasoning community.

We (my thesis advisor and I) are currently exploring how the representation of knowledge goals can be used to provide *guidance* in an interactive system. To that end, we're currently working on an idea of a defined goal space via a vector representation of knowledge goals, thereby allowing us to ascribe distances to knowledge goals. In the context of an interactive dialogue, a goal trajectory can be specified by a distance, e.g., the total distance from one goal to another; a system can therefore be said to provide guidance by minimizing or accelerating the path to a goal or set of knowledge goals. We are currently conducting an empirical study to confirm this hypothesis, possibly presented as a student poster to AAAI 2016 and prepare a demo that will be ready in time for the Doctoral Consortium.

References

Aha, D. W.; McSherry, D.; and Yang, Q. 2005. Advances in conversational case-based reasoning. *The knowledge engineering review* 20(03):247–254.

Bengfort, B., and Cox, M. T. 2015. Interactive Knowledge-Goal Reasoning. In *Goal Reasoning: Papers from the ACS Workshop*, 10.

Berant, J.; Chou, A.; Frostig, R.; and Liang, P. 2013. Semantic Parsing on Freebase from Question-Answer Pairs. In *EMNLP*, 1533–1544.

Cox, M. T., and Veloso, M. M. 1998. Goal transformations in continuous planning. In *Proceedings of the 1998 AAAI fall symposium on distributed continual planning*, 23–30.

Cox, M. T., and Zhang, C. 2007. Mixed-initiative goal manipulation. *AI Magazine* 28(2):62.

Kolodner, J. 1993. *Case-based reasoning*. Morgan Kaufmann.

Lopez De Mantaras, R.; McSherry, D.; Bridge, D.; Leake, D.; Smyth, B.; Craw, S.; Faltings, B.; Maher, M. L.; Cox, M. T.; Forbus, K.; and others. 2005. Retrieval, reuse, revision and retention in case-based reasoning. *The Knowledge Engineering Review* 20(03):215–240.

Ram, A., and Hunter, L. 1991. A Goal-Based Approach to Intelligent Information Retrieval. In *Machine Learning: Proceedings of the Eighth International Conference*, 265–269. Citeseer.

Ram, A. 1991. A theory of questions and question asking. *Journal of the Learning Sciences* 1(3-4):273–318.

Wang, G.; Gill, K.; Mohanlal, M.; Zheng, H.; and Zhao, B. Y. 2013. Wisdom in the social crowd: an analysis of quora. In *Proceedings of the 22nd international conference on World Wide Web*, 1341–1352. International World Wide Web Conferences Steering Committee.

Yahya, M.; Berberich, K.; Elbassuoni, S.; Ramanath, M.; Tresp, V.; and Weikum, G. 2012. Natural language questions for the web of data. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, 379–390. Association for Computational Linguistics.

Benjamin Bengfort

Curriculum Vitae

73 Bryant St NW
Washington D.C. 20001
(701) 680 3095
benjamin@bengfort.com
cs.umd.edu/~bengfort

Education

2014–present **PhD Candidate**, *University of Maryland, College Park*.
Graduate Research Assistant, Metacognitive Lab

2008-2010 **M.S. Computer Science**, *North Dakota State University*.
Phi Kappa Phi, Upsilon Pi Epsilon

2004-2006 **B.A. Economics**, *University of Maryland, College Park*.
Primanum Honor Society

2002-2003 **English Major**, *United States Naval Academy*.

Publications

Benjamin Bengfort and Michael T. Cox. Interactive Knowledge-Goal Reasoning. In *Goal Reasoning: Papers from the ACS Workshop*, page 10, 2015.

Benjamin Bengfort, Philip Y. Kim, Kevin Harrison, and James A. Reggia. Evolutionary design of self-organizing particle systems for collective problem solving. In *Swarm Intelligence (SIS), 2014 IEEE Symposium on*, pages 1–8. IEEE, 2014.

Benjamin Bengfort and Xiaojiang Du. Efficient resource allocation in Hybrid Wireless Networks. In *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*, pages 820–825. IEEE, 2011.

Benjamin Bengfort and Jenny Kim. Hadoop Fundamentals for Data Scientists, January 2015.

Tony Ojeda, Sean Patrick Murphy, Benjamin Bengfort, and Abhijit Dasgupta. *Practical Data Science Cookbook*. Packt Publishing Ltd, 2014.

Experience

2014–present **Adjunct Faculty**, *Georgetown University*, Washington, DC.
Lecturer, Data Analytics Certificate Program (CCPE). Teaching *Foundations of Data Science*, *Software Engineering for Data*, *Data Sources*, *Data Analysis II: Machine Learning*, and *Applied Data Analytics*.

2013–2014 **Data Scientist**, *Cobrain Company*, Bethesda, MD.
Developed a Global Recommendation Engine, using Collaborative Filtering algorithms as well as Active Learning and adaptive systems from a machine-learning standpoint. Graph traversal and clustering across massive data stores required distributed Graph databases: Titan, as well as strong computation in Hadoop and Python with Pandas and NumPy. Entity resolution and ontological reasoning using probabilistic graphical models and RDF ontologies also factored into the work to disambiguate crawl data from a massive distributed crawler also developed at Cobrain.

2011–2013 **Chief Technology Officer**, *Unbound Concepts*, Baltimore, MD.
Natural Language Processing across a large dataset of children's literature using NLTK
Machine Learning and predictive analysis with clustering and multivariate linear regression
Big data analysis applied to the NLP and ML using Hadoop and MapReduce techniques.Mobile Application development: Android, iOS: Developed the BookLeveler app
Rich web application development with Django along with JSON server for mobile apps.

2010–2012 **Lead Programmer**, *Tactical Network Solutions*, Columbia, MD.
Python software development for embedded, mobile and server applications
Large-scale and real-time data analysis of Petabytes of wireless collects.
Wireless offensive computer network attack, computer network operations
Real time asset tracking software development with geolocation and KML toolkits
Django web application development and Android mobile client applications development

2007–2009 **CMS Analyst**, *Oxford University Press*, Oxford, United Kingdom.
Management of custom content management solutions for electronic publishing division
Customer representative for agile based development of OED 4.0 CD-ROM software
Data and database management for XML DTDs and Schemas

2003–2004 **Junior Network Engineer**, *CenGen, Inc.*, Columbia, MD.
Network support for DARPA's Grand Challenge & USMC Condor
Mobile and wireless integration, network vehicle integration
Field support and logistics, small team leadership

Skills

Software Development	Python 2.7 & 3.4, GoLang, JavaScript, Java, C/C++, Git, Agile	Big Data	Spark, Hadoop, MapReduce, HDFS, Distributed Systems, Celery
Machine Learning	Scikit-Learn, SparkML, GraphX	NLP	NLTK, TextBlob, Pattern
Application Development	Django, Nginx, AJAX, jQuery, Underscore, Bootstrap, Flask, Django-Rest-Framework, RESTful APIs	Databases	PostgreSQL, MongoDB, LevelDB, Redis, Hive, Titan, Neo4j, MySQL, SQLite
Networking	Kismet, Wireshark, Mgen, 802.11, Protocol Buffers, NS2	OS	Mac OS X, Ubuntu Linux, Gentoo, Embedded Ubuntu, Android, iOS.

Volunteer Experience

- Advances in Cognitive Systems 2015: Referee and reviewer of paper submissions.
- Data Community DC: Treasurer, Board Member, and Organizer for DIDC.
- Vino Scholastico: Charity committee for a tuition fundraising event at HCC in Maryland.
- Team Fight, The Ulman Cancer Foundation: athletes raising money for the long term cancer treatment and care of young adults.
- The Bridging Project: taught level 2 maths to asylum seeking refugees.
- Primannum Honor Society: organized blood drives for The Red Cross.
- Leadership U: tutored ESL elementary aged students.
- National Honor Society

Personal Statement

Benjamin Bengfort *bengfort@cs.umd.edu*

September 22, 2015

My graduate academic advisors have had extremely good fortune upon accepting me as their student; three of them were promoted to take larger roles at more prestigious universities, and one was offered a lucrative position at Bell Labs. Obviously many factors led to their success, but since 100% of my advisors have had promotions within a year of advising me I suspect that I exist in some sort of tenure projection field that has aligned them well with their respective search committees. Or it could be that I simply have good taste in advisors. In either case, the departure of my advisors has led to a rather tumultuous academic career, observable in the wide breadth of topics of my academic publications.

It is perhaps for this reason that I have a hard time answering the question, "what year are you?" I am either in my third year or my eighth, depending on who you ask and the mood of the Graduate School. But what I can't deny, is that due to gaps, fits, and starts in my pursuit of the academic dream, I do have a more comprehensive work experience than some of my fellow, younger classmates. As a result, I'm much more comfortable answering the question "why do you want your PhD?" than they are, particularly as my graduate experience has not been to improve my opportunity to acquire a job (nor in the pursuit of a University position), but rather is an investment in myself so that I can pursue something that will change the world.

My professional career has been one of startups, large and small. Being in a startup is hard, it requires commitment, dedication, and a little bit of self-sacrifice. But the entrepreneurial spirit of taking risks to do something big, often makes those difficulties worth it. Although the fixtures are changed, pursuing a PhD requires that same entrepreneurial character, a willingness to be a jack-of-all-trades, to do what needs to be done if only to make a small contribution to human knowledge. I believe that the disruptive influence of startup culture, combined with my varied academic background has made me able to think and see new opportunities, integrations, and cross-disciplinary approaches that I have brought to my research at the University of Maryland.

Now at last I'm approaching the final stretch of my academic career. This semester I will have completed all of the course and residency requirements at the University of Maryland, where I have finally landed. I have a committee committed to my work, promotions aside, and I am preparing to propose a dissertation in the field of Artificial Intelligence at the end of next semester. Although a diverse background has brought me this far, it is time to focus towards an in-depth study where I can contribute my research; particularly a study of computational investigation and the trajectories of knowledge goals.

By participating in and presenting at the AAAI Doctoral Consortium, I hope to refine my study and research interests to a laser focus and to complement the advice of my academic advisor. I believe that the DC will be a critical step towards understanding my topic in the context of Artificial Intelligence research, and that the contextualization of my research will fine tune the path to the completion of my dissertation. I have not participated in any previous doctoral consortia, and I am not applying to any others in 2016 because I believe the AAAI DC is the correct place to present my research and the only one that will meet this goal. In return, I hope to bring my startup and entrepreneurial energy to the DC as an active and lively participant; encouraging my peers and engaging with more seasoned researchers. And perhaps we may even observe more effects of the tenure projection field!